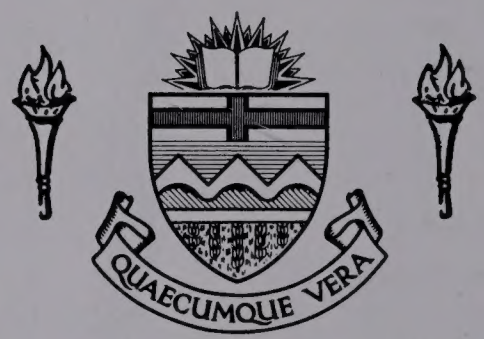


The University of Alberta
Printing Department
Edmonton, Alberta

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR . . . . . JAMES AMOS ORIEUX . . . . . .

TITLE OF THESIS . Queensland Test and Coloured Progressive . . . .
. . Matrices: A Normative Study on Deaf Children

DEGREE FOR WHICH THESIS WAS PRESENTED .Master of.Education. . . .

YEAR THIS DEGREE GRANTED . . . . Fall, 1974. . . . . .

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

DATED . Sept.

THE UNIVERSITY OF ALBERTA

QUEENSLAND TEST AND COLOURED PROGRESSIVE MATRICES: A NORMATIVE STUDY ON DEAF CHILDREN

by

© James A. Orieux

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

IN

COUNSELLING PSYCHOLOGY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1974

THE UNIVERSITY OF ALBERTA

QUESTIONED TEST AND COUNCIL PRODUCTIVE MATERIALS: A NORMATIVE STUDY ON DEAF CHILDREN

BY

Jenna A. Delaune

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF EDUCATION IN COUNSELLING PSYCHOLOGY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1974

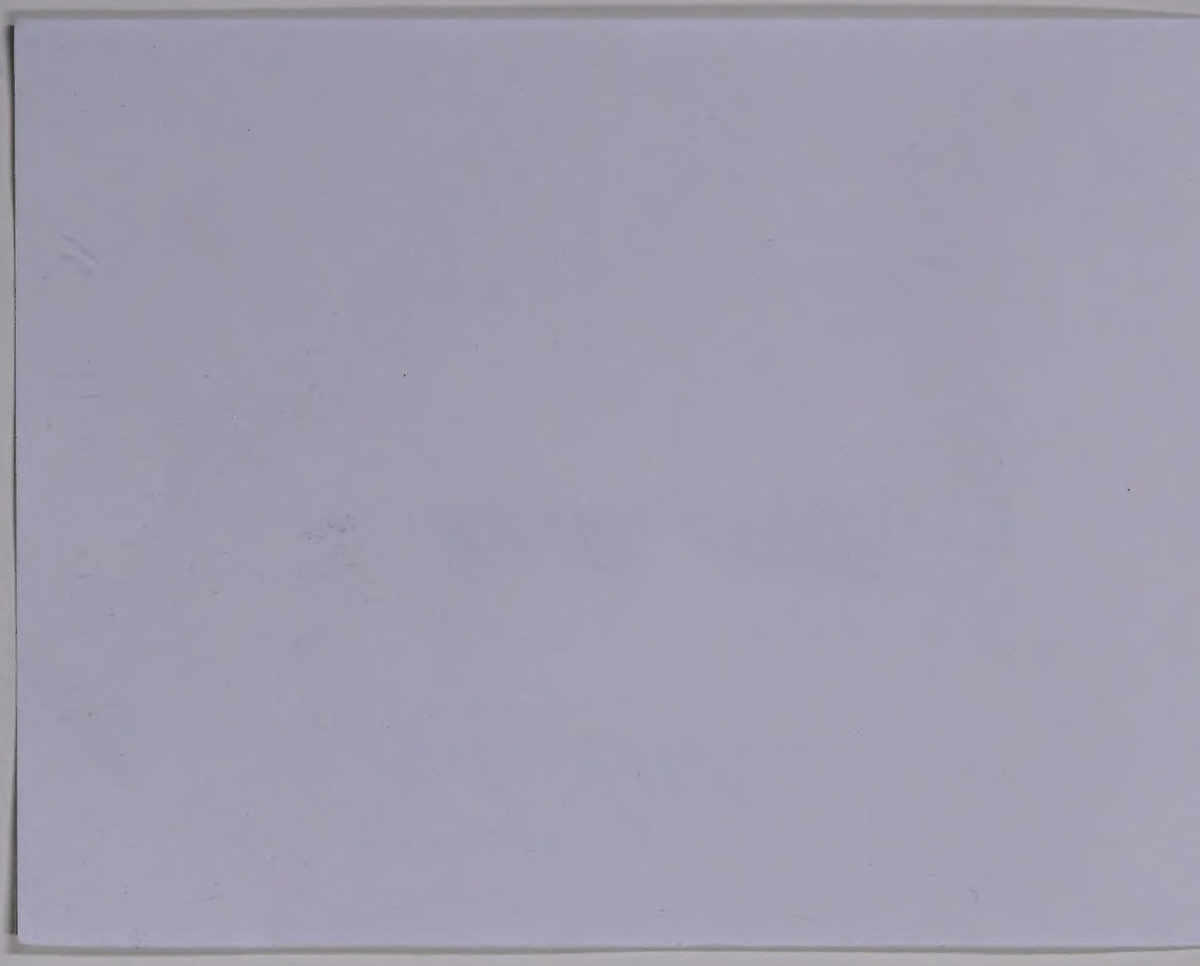
THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Queensland Test and Coloured Progressive Matrices: A Normative Study on Deaf Children", submitted by James A. Orieux in partial fulfilment of the requirements for the degree of Master of Education in Counselling Psychology.

Supervisor

Date Sept. 12, 1977



ABSTRACT

The primary purpose of this study was to develop preliminary norms for the students at the Alberta School for the Deaf for the Queensland Test, which is a test of non-verbal cognitive ability.

Ninety-four students, 54 boys and 40 girls, at the Alberta School for the Deaf were individually administered the Queensland Test and the Raven's Coloured Progressive Matrices (1947), which was used as a comparative measure of non-verbal cognitive ability. The latter was administered as a group test. A significant correlation of .80 ($p<.001$) was obtained between the Queensland Test and the Coloured Progressive Matrices.

The relationship between the effects of sex, residence status (time spent away from school residence), and hearing level, and students' performance on the Queensland Test was investigated using analysis of variance. Results were not significant at the .05 level.

Preliminary norms were developed and reported as Quartiles. The small number of cases at each age level precluded the development of more refined norms such as percentiles or stanines. Guidelines were provided for setting up more extensive norms.

ACKNOWLEDGMENTS

This research project was made possible through the assistance and cooperation of many people. I wish to thank, in particular, Dr. Peter Calder for his patience, direction and encouragement; Dr. Russell MacArthur for his personal interest and his assistance, especially in the area of statistics and norms; Dr. Oliva Gironella for her enthusiasm, cooperation and assistance in the collection of the data at the School for the Deaf; and Mrs. Mary Reeve for her skillful typing.

I would also like to thank the Principal, Staff and students at the Alberta School for the Deaf for their participation in the project, and Mrs. Sonia Hayward for sharing her knowledge and contributing some of the data.

Then singular thanks to my wife Bev and my daughter Anne for their faithful support, assistance and remarkable patience.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Those working with the handicapped, and specifically with the deaf, are often frustrated by the lack of assessment instruments available which have been developed for and normed on a deaf population. Levine (1971) states that in thinking out the key problem in testing the deaf, it occurred to her that ' ... there are not enough instruments devised for or standardized on the deaf to cover the many needs and situations that call for valid test information'. She goes on to say, 'As a result of this impoverishment, testers of the deaf supplement their slim resources by dipping into the pool of measures standardized on the hearing (p. 83)'. Kearney (1969) states

> The assumption can be made that, for the deaf, a performance scale of cognitive ability may be able to give a useful assessment of mental ability, whereas more frequently used scales which contain verbal items, may give a less valid assessment. In other words, an intelligent deaf subject who is rated low on a verbal intelligence test, is rated low because of his deafness. A wrong diagnosis of his intelligence has been made (p. 3).

She sees two main problems confronting testers of the deaf: first, finding a scale of cognitive ability which does not require the use of overt language; and second, establishing that the test items do not require the use of 'internal language' for the solution of the set task. In this case, she defines "internal language" as ' ... the use of words as opposed to the use of symbols in the forming of concepts and ideas (p. 4)'.

In her study, in which she used the Queensland Test with

deaf children (Parker, 1965; Kearney, 1969), Kearney concludes that ' ... The Queensland Test is a satisfactory measure of the cognitive ability of deaf children in Queensland'. She goes on to say

> In order ... to assess the ability of the deaf to be assured that this is developed to the fullest possible extent of the potential that exists, there is a need for a measuring tool that will accurately determine cognitive capacity. The Queensland Test has been shown to be an adequate test in its present form. It remains that norms should be developed for the deaf population (p. 13).

The Problem

It appears that there is a need for tests which are developed for and/or standardized on a deaf population. The concept of intelligence in the general population is strongly associated and dependent upon verbal ability and the related tasks of reading and comprehension. Although there are many excellent non-verbal "performance scales" of cognitive ability, the manner in which the procedure for administering the items are standardized may often be biased toward a hearing child, especially in the case of timed items on which a hearing child can cue himself verbally in order to retain concepts such as colors and numbers. Items dependent upon a time limit may present undue difficulty to the deaf child. The deaf child's system of symbols may be quite different. He may not have the same facility of utilizing visual cues in forming concepts and, according to Vernon and Brown (1964) the deaf child reacts to the factor of timing on a test, differently from the hearing child.

In the assessment of deaf children, there may be valid reasons for wanting to compare the deaf child to a hearing child, or the norms of the hearing world, in which he must eventually cope. However, there

are many occasions, especially in clinical assessment, when a comparison of the deaf child with his peers is very valuable, especially in a residential school setting where conditions are controlled to a greater degree than in a regular public school. Well known tests, such as the Wechsler Intelligence Scale for Children - Performance Scale, (Wechsler, 1949), can be used effectively in assessing the deaf on an individual clinical basis. However, there are no deaf norms available and it appears that modifications to certain subtests would have to be made before such norms can be developed (Graham & Shapiro, 1953; Levine, 1971; Willis, Wright & Wolfe, 1972).

Purpose of the Study

The primary purpose of this thesis is to establish preliminary norms to be used with deaf children enrolled at the Alberta School for the Deaf in Edmonton, for the Queensland Test. It is felt that this test can be administered without modification and that it will provide a useful measure of cognitive ability. A secondary purpose is to investigate the usefulness of the test with respect to the validity and reliability reported in Kearney's study (1969). Performance on the Queensland Test will be compared to performance on the Coloured Progressive Matrices (Raven, 1947), a non-verbal intelligence test. The effects of sex, level of hearing, and residence status on a subject's performance on the Queensland Test will be investigated.

Definitions

1. Cognitive Ability - non-verbal intelligence as measured by the Queensland Test and the Coloured Progressive Matrices.

2. Hearing Level - as contained in the School's audiological reports. The figure reported is the average decibel loss for the better ear within the "sppech awareness threshold", that is, the level at which the subject is aware of sounds resulting from speech, but not necessarily the ability to discriminate speech sounds in terms of comprehension. Decibel loss is generally reported as a figure from 0 - 100 with an increase in the figure indicating a greater hearing loss.

3. Normative Group - Younger students are admitted to the Alberta School for the Deaf on the basis of having been deaf at birth or prior to having acquired language skills. Those students admitted at an older age are accepted on the basis of not being able to cope in a regular school program. A child who through assessment is diagnosed as mentally retarded, is not admitted to the School.

The normative group consisted of 94 students (54 boys and 40 girls) between the ages of 7 years, 0 months and 15 years, 0 months as of April 1, 1974. The range in age was selected on the basis that the test being used is not recommended for use below age 7 years, and the top age of 15 years, represented a practical cut-off point in terms of cognitive ability and future use of the test in assessing students at the School. All students within the prescribed age limit were tested, regardless of whether they were receiving special programs relevant to learning difficulties.

Limitations

1. The results of this study were intended to be applied to those students enrolled at the Alberta School for the Deaf and their use with other deaf children should be done with great care.

2. It is recognized that there are other performance scales of intelligence, presently based on the hearing, which could be slightly modified and used for the same purpose of establishing norms on a deaf population. The Queensland Test was selected for its facility of administration under conditions of reduced communication, making modifications unnecessary; a simple system of scoring; the fact that the five subtests of the scale are tasks which are presently used on existing performance scales for both hearing and deaf populations; and finally, the fact that the Test is presently being used at the Alberta School for the Deaf on an individual clinical basis, but without the use of comparative data from which to draw inferences.

3. The reporting of normative data in this study was limited by the small number of cases in any one age category. However, it was hoped that this project would initiate the building up of norms in the future, and that with increased use, the strengths and limitations of the Queensland Test when used with deaf children would become more evident.

CHAPTER II

RELATED RESEARCH AND LITERATURE

Conflicting results have emerged in research dealing with the intelligence of the deaf which in a sense appear to be test dependent. Goetzinger, Wills and Dekker (1967) state that many of the large scale studies of the past found the deaf to be retarded in intelligence from 2 to 3 years, specifically when either the Pintner Non-Language Test or the Digit-Symbol Test was used, and cite relevant studies: (Pintner & Patterson, 1918; Reamer, 1921; Day, Fusfeld & Pintner, 1928; MacKane, 1932). They also refer to Zeckel and Van der Kolk (1939) who reported the deaf to be retarded on the Porteus Maze, and to Oleron (1950) who found retardation of about 2 years on the Raven's Progressive Matrices, 1938.

Contrary evidence is put forth by Goetzinger and Rousey (1957):

> Drever and Collins (1936)... administered their own performance test of intelligence to both deaf and hearing subjects, and found the former to be equal or slightly better than the latter... Subsequent research (Myklebust & Burchard, 1945; Lyons, 1933; MacKane, 1932) has shown the deaf to test within normal limits intellectually when performance scales are used (p.388).

Myklebust (1964) states 'Psychologists continued to study the possibility that the child with severe hearing loss was mentally retarded (.67)'. He cites studies (Schick, 1934; Streng & Kirk, 1938; and Myklebust and Burchard, 1945) which indicated that when individual performance tests were used, children in the schools for the deaf were of average intelligence, and which contradicted the findings of Pintner who used group tests to facilitate the obtaining of larger samples.

The question of whether the intelligence of the deaf is comparable to that of the hearing has led to a fair amount of research pertaining to the psychology of the deaf. Furth (1964, 1966, 1973) stands out as a leading figure in research dealing with the thinking of the deaf and with the relationship of intelligence and non-verbal thought processes. He addresses himself to the question of interpreting the cognitive ability of the deaf as compared to the hearing, and is prompted to inquire '... whether the differences observed in deaf people's behavior are traceable to a different cognitive structure which they develop because they lack language. How should one describe a cognitive structure? To what cause should one ascribe deficiencies which make the deaf appear "concrete-minded" (Furth 1966, p.2)?' Furth then goes on to dispel the assumptions commonly held about the relationship of language and thinking, such as '... the theory that language faithfully mirrors thinking; that language is almost the only important symbol system, and that language, symbols, and thinking are necessarily, even inseparably linked (p.4)'. He points out that the terms "concrete" and "abstract" were used in rather loose fashion and verbal behavior was identified with abstract or conceptual behavior while nonverbal behavior was associated with concrete behavior on a perceptual level. From his own experience with deaf adults and from some preliminary first findings, Furth thinks that it seems more scientific to start with the hypothesis '... namely, to expect no differences in cognitive structure between the hearing and the deaf, rather than approach with the assumption that a

necessary relation between language and thinking does exist (p.4)'. He concludes 'Thus, those who have hypothesized differences in the thinking of the deaf usually employed verbal tests and found their theory borne out as they observed that deaf persons were in general poorer than hearing persons on such tasks (p.5)'.

The assessment of the deaf with a variety of performance scales and nonverbal tests, some of which have been normed on the deaf (Hiskey, 1966; Snijders-Oomen, 1958), has also opened up an area of conflicting research, but, nevertheless contributed much worthwhile information in the comparison of the deaf with their own peers and with the hearing.

Levine (1971) provides a comprehensive chronological listing, comprised of 36 studies dating from 1930 to 1970, pertaining to the evaluation of mental tests used with the deaf.

There is concern expressed with regard to the use of standardized tests normed on a hearing population in the assessment of deaf subjects. Furth (1966) states

> The first consideration in testing deaf persons is a very obvious one. Clearly, verbal tasks cannot be relied upon as accurate measurements of intelligence in language-deficient subjects. While not denying that within certain limits, when used and interpreted correctly, such tests may give some indication of a deaf person's performance level, we believe most standardized intelligence tests, relying heavily on verbal ability, would prove very inaccurate indices of cognitive ability. Also, any standardized intelligence test depends for its validity on the assumption that a given subject is part of the sample on which the test was standardized. The deaf, of course, are not part of this sample (p.71).

Levine (1971) states 'There are simply not enough instruments devised for or standardized on the deaf to cover the many needs and situations that call for valid test information ... as a result of this impoverishment, testers of the deaf supplement their slim resources by dipping into the pool of measures standardized on the hearing (p.83)'. This same concern is reflected in several other studies (Hiskey, 1956; Birch, 1958; Myklebust, 1964; Furth, 1966, 1973; Kearney, 1969; Hanson, Hancock & Kopra, 1969; Willis, Wright & Wolf, 1972).

There is the opinion that since the deaf are, for the most part, being educated to cope in the normal world, they should be assessed and compared to the hearing. Graham and Shapiro (1953) investigated the use of the Wechsler Intelligence Scale for Children (WISC) with deaf children on the premise that '... much of the emphasis in the education of the deaf has been in preparing them vocationally so they may be self-supporting', and hence '... since in most instances the deaf must compete with the hearing, it seems reasonable that tests for both groups be the same, modified only when necessary to exclude items dependent upon hearing (p.396)'. They equated 20 deaf subjects on the basis of intelligence, sex, nativity, age, color, and health, with normally hearing subjects, in two groups of 20 each. The deaf as well as one of the hearing groups were given the WISC using pantomime directions. The other group of hearing children took the test under standard procedure. The mean IQ's for the deaf and the hearing group with pantomime directions were respectively 96.1 and 95.5. The mean IQ for the hearing group which took the test when verbal or standard

directions were employed, was 101.15. Graham and Shapiro therefore suggested that a correction factor of 5 points be applied when the WISC is utilized with deaf subjects. They observed that the pantomime directions had an effect on the group of hearing children and, among other findings, that the deaf subjects were affected by the time factor on some of the subtests. They concluded that since the effect of pantomime must be considered in any test to be used with the deaf, the test could not be used validly to obtain a measure of intelligence for deaf children. They then suggested a correction factor to nullify the effects of pantomime instructions and stated '... as it is, the test can be administered via pantomime as a crude measure (p.398)'.

Evans (1966) compared 100 deaf and partially hearing children on the WISC (Performance) and Raven's Standard Progressive Matrices. Three years later, 42 of the subjects were retested. The WISC (Perf.) results had high internal consistency, .98 and .89 for age groups 6-10 years and 11-15 years respectively, high retest reliability (.98), and were distributed within reasonably normal limits. The Progressive Matrices test was also a highly reliable measure of current intellectual output, .90 and .92 respectively for the same age groups. However, the results were substantially below normal and showed some fluctuations with time. There was low intercorrelations with the WISC (Perf.) results. Evans, in discussing the poor performance of the deaf subjects on the Progressive Matrices, puts forth the hypothesis that linguistic deprivation due to deafness might impede the development of abstract reasoning, and relates this to some of the subtest results on the WISC: subjects did

well on the Block Design and Picture Completion (concrete ability) and performed poorly in Coding and Picture Arrangement (more abstract and similar to Progressive Matrices). Graham and Shapiro (1953) obtained similar findings on the four subtests mentioned. Evans concludes that the WISC is a valid measure of the general ability of deaf children.

Willis, Wright and Wolfe (1972) compared 40 deaf and 40 hearing children on the WISC (Perf.) and the Nebraska Test of Learning Aptitude (Perf.) (Hiskey, 1955), and found no significant differences between deaf and hearing subjects on the WISC performance IQ. Lower scores were found on the Picture Arrangement subtest, but the deaf children as a group did not appear to respond significantly different from hearing children to visual tasks - Picture Completion (WISC) and Completion of Drawing (Nebraska). They found low correlations between the WISC and the Nebraska for deaf subjects and state that this may be explained on the basis that, unlike the Nebraska, the WISC has never been standardized separately for deaf subjects, thus there are no separate norms.

Vernon and Brown (1964) discussing psychological tests and testing procedures in the evaluation of deaf and hard of hearing children, point out the caution necessary in selecting tests and the fact that not all non-verbal or performance tests are suitable for use with the deaf in that they may require some verbal instructions and may also contain a time factor. They support one of the findings of Graham and Shapiro (1953) pertaining to timed items on tests used with the deaf:

> Intelligence tests for young deaf or hard of hearing children (age 12 and below) that emphasize time are not as valid in most cases as are other tests which do not stress time. This is because these children often react to the factor of timing by either working in great haste and ignoring accuracy or else

> disregarding the time factor completely. In either case, the result is not necessarily a reflection of intelligence. (p.418).

Hiskey (1956), in comparing hearing and deaf children on the Nebraska Test of Learning Ability and specifically on the Memory for Colored Sticks item, found that when only one color was demanded the performances of the two groups were quite close. However, when two or more colors were presented the hearing children used their speech and hearing to an advantage. They would repeat the colors aloud while searching for the proper sticks, and this apparently strengthened their retention. The same appeared to be true on the Memory for Digits item. Hiskey concludes 'It would appear that even a non-verbal response of the immediate recall type often is strengthened through the utilization of supplementary sensory and motor cues (p.332)'.

In dealing with the conflicting view of whether to compare the deaf to the hearing or to a deaf population, in the use of psychological tests, Levine (1971) sees the problem as one of semantics in the use of the word 'world' when it is said that the deaf live in the same world as the hearing:

> As used in psychology, and specifically in test construction, the term "world" does not refer to a place but to a body of experiences and, in fact, to the specific bodies of experiences on which tests are standardized.

She sees the deaf child's world as

> ... a place of widely differing sensory, perceptual, and associated life experiences from those of the hearing.

Consequently;

> ... even though deaf and hearing children live in the same place - even the same home - they can nevertheless be worlds apart psychologically (p.83).

Levine cites Birch (1958) as holding the view that

> ... in the psychological appraisal of seriously handicapped children, the major purpose and immediate concern are to determine the kinds of training and experience that will best promote the child's functional adaptive abilities rather than to predict whether the child will eventually be able to compete successfully with his age-peers from the population at large (p.84).

Kearney (1969) the only one, to date, who has conducted a study on the use of the Queensland Test with deaf children, saw a need for an investigation into the availability of tests of the general cognitive capacity of the deaf so that the serious implications of deafness being associated with their cognitive functioning could be determined. Her research was undertaken in order to investigate more fully the adequacy of the Queensland Test as a measure of the general cognitive ability of deaf subjects. She recognized the fact that an adequate test for the deaf should meet the following conditions:

1. not require the use of overt language in giving the instructions or in order to communicate the problem;
2. the responses must not be oral or written, that is, they should be such as can be visually demonstrated and imitated;
3. the test items should not require 'internal language', that is the use of words as opposed to the use of images in the forming of concepts and ideas, for the solution of the set task. The Queensland Test met the above criteria (p.4).

In discussing the construct validity of the Queensland Test, Kearney refers to theorists who have studied human intelligence (Vernon, 1965; Thurstone, 1957; Guilford, 1957; Wechsler, 1958) and states that

'... for the purpose of this research, the problem of the exact nature of intelligence must be postponed in favour of the investigation into the measurement of cognitive ability', and then states

> It would seem that the five subtests of the Queensland Test involve the measurement of memory (retention), motor skills (reproduction), cognition, observation (visualization), reasoning, abstraction, spatial and number skills. None of these is measured by fewer than three of the sub-tests nor more than five. Each of the tests in turn measures at least three of these abilities (p.11).

She recognizes that verbal factors such as Thurston's 'word fluency'. Wechsler's 'verbal comprehension' and Vernon's 'v:ed', are not measured by the Queensland Test and says 'The issue that remains is whether intelligence can be measured without taking a verbal factor into account (p.11)'. However, she points out that it has already been shown that the deaf cannot be fairly assessed using tests which contain verbal items and that the non-verbal tasks must be divested of verbal components even when these are in the form of internal verbalization.

In her study which was conducted in 1965-66 (Kearney, 1969) 120 children from the School for the Deaf in Brisbane, aged 7-14 years, were matched with a similar number of normal children in the same age and grade level, from a typical state school, and administered the Queensland Test. One year later, the eight year old population (15 Ss) were retested using a wide battery of tests including the Queensland Test. Briefly, the results were:

1. There was a significant difference between the deaf and hearing samples ($p<.001$);

2. Test-retest reliability: using Queensland Test (QT) total scores, the results were compared and indicated a high test-retest reliability ($r = .82$) for this small sample;

3. QT scores were correlated with teachers' ranking on academic achievement, using Spearman's rank order correlation. The results (P = .69) indicated the Queensland Test compares favorably with academic achievement and general classroom ability;

4. QT scores were converted to tentative IQ norms and compared with the WISC and the Non-veral Intelligence Tests for Deaf and Hearing Subjects or S.O.N. (Snijders-Oomen, 1959). The following correlations were obtained: QT and WISC, r = .825; QT and S.O.N., r = .718; WISC and S.O.N., r = .749.

Those subjects who performed better on the Queensland Test than on the other two tests were all poor verbalizers and had 'very severe' to 'profound' hearing losses. The converse was true: those who performed less well on the Queensland Test than on the S.O.N. and WISC were all good verbalizers and had less severe hearing losses. It was concluded that their verbal ability may have influenced their scores on the WISC and S.O.N. As well, it would seem that the Queensland Test may be a better measure for those who are not so proficient verbally.

One additional study was found which made use of the Queensland Test, but with a subnormal population, namely the mentally retarded (Cartan, 1971). This study was not available in print. However, a few observations from the abstract are worth mentioning. The Queensland Test was used with both children and adults. Correlations were obtained between the QT and the WISC, WAIS, and the Stanford-Binet, Form L-M. The correlations are not reported but it is mentioned that the highest correlations were obtained between the QT and the WISC full scale and WISC performance scale. Correlations between test scores and teacher ratings were also carried out and it was observed that the Queensland Test's usefulness with the population studied may be mostly in predicting classroom performance.

The Progressive Matrices Test (Raven, 1938, 1947) has been used to some extent in the assessment of the deaf, with conflicting results.

Oleron (1950) tested 246 deaf subjects between the ages of 9 and 21, using the Progressive Matrices (1938). Results indicated the deaf to be inferior and with slower mental development. He concluded that 'Deafness impairs capacity to use abstract forms of intellectual activity (p.192)'. Vernon and Brown (1964) rated the Progressive Matrices '... good as a second test to substantiate another more comprehensive intelligence test ... easy to administer and score ... yields invalid test scores for deaf children who are impulsive and who tend to respond randomly rather than with accuracy and care (p.416)'.

As reported earlier in this chapter, Evans (1966) indicated the Progressive Matrices test to be a highly reliable measure of current intellectual output but that the results were substantially below normal, showed some fluctuation with time and correlated quite low with the WISC performance scale. Goetzinger, Willis, and Dekker (1967) tested 96 deaf subjects on the Progressive Matrices, 1938. Results indicated them to be approximately 2 years below the norm for comparable ages.

Goetzinger and Houchins (1969) conducted a study using the 1947 Coloured Progressive Matrices with deaf and hearing subjects. They also investigated methods of directions: pantomime compared with verbal, and explored the influence of sex on test performance. The sample was comprised of 40 deaf and 40 hearing children equally divided into two

age groups, 6½ and 8½ years of age. Results indicated no significant difference as a function of age level or sex; no difference in age between the deaf and hearing except between older and younger subjects, which was to be expected; both deaf and hearing improved in performance as a function of age; pantomime directions did not significantly alter performance; both groups were within the normal range of intelligence as measured by the Nebraska Test of Learning Aptitude and were not significantly different on the Progressive Matrices; elementary school deaf children performed as well as their hearing counterparts on the 1947 Coloured Progressive Matrices. Goetzinger and Houchings reported the Progressive Matrices was easy to administer and convenient in that it does not require the services of a school psychologist for administration.

Establishing Norms.

The difficulties inherent in establishing age norms with small numbers for each age group are presented and discussed in most texts dealing with statistics and educational measurement. However, guidelines in setting up preliminary norms, as in this thesis, are not offered (Ferguson, 1971; McNemar, 1969; Guilford, 1973; Thorndike & Hagen, 1969). The Standards for Educational and Psychological Tests and Manuals (American Psychological Association, 1966) provides some parameters within which one can conduct a project of a preliminary nature such as the one being presented in this thesis. Following are some pertinent sections:

F5. Norms should be reported in the test manual in terms of standard scores or percentile ranks which reflect the distribution of scores in an appropriate reference group or groups. ESSENTIAL (p.35).

F6. Norms presented in the test manual should refer to defined and clearly described populations. These populations should be the groups to whom users of the test will ordinarily wish to compare the persons tested. ESSENTIAL (p.35).

F6.31. If the sample on which the norms are based is small or otherwise undependable, the user should be cautioned explicitly in the test manual regarding the possible magnitude of errors arising in the interpretation of scores. ESSENTIAL (p.36).

Bean (1953) states

> If the individual who constructs the test is unable to obtain a very large group that is in any sense "typical", he may use deciles instead of 100 to express a rough position for scores in a smaller population. If the resulting curve from a group of limited size seems to approach a normal distribution fairly closely, this finding at least indicates that the test is suitable for that particular group ... Norms from small samples should be used with caution (p.174).

To summarize, the literature reviewed provided an overview of research concerned with the intelligence of the deaf relative to the hearing world and, more specifically, to the assessment of the intellectual ability of the deaf and the instruments employed in the process. Research pertaining to the Queensland Test and the Progressive Matrices, the instruments used in this thesis, was also presented. Finally, a brief reference to the establishment of norms based on a small sample was put forth.

CHAPTER III

PROCEDURE AND DESIGN

Introduction

The primary purpose of the study was to develop norms for the Queensland Test based on a population of deaf children attending a residential school for the deaf.

Performance on the Queensland Test was compared to performance on the Progressive Matrices, which also yields a measure of non-verbal intelligence or cognitive ability. The relationship of the variables of sex, level of hearing, and residential status with the Queensland Test was observed and the results reported.

INSTRUMENTS

Queensland Test

Developed at the University of Queensland and published in 1970, the Queensland Test is based on cross-cultural research conducted over a period of ten years and sponsored primarily by the University of Queensland and the Department of Health, Education and Welfare. The major preliminary studies were conducted among indigenes of Papua-New Guinea and the Aborogines of Australia, with comparative data from Australians of European descent. One of the latter samples was the subject of a study conducted on 120 pupils from the school for the deaf in Brisbane (Kearney, 1969). The Test is essentially based on the unpublished PIR (Pacific Island Regiment) Test and the New Guinea Performance Scale (See McElwain & Kearney, 1970).

The QT is defined as a performance test of general cognitive capacity designed for use under conditions of reduced communications, and is administered completely without verbal behavior by either the examiner or the subject. It consists of a battery of five subtests of varying kinds:

1. Knox Cubes - consists of four black 7/8 inch cubes spaced at equal intervals and attached to a white base frame 8 inches long. There are two detached black 7/8 inch cubes used for tapping. The examiner gives the subject a detached cube and with the other in his own hand taps a sequence on the frame and then invites the subject to tap the same sequence on the frame;

2. Beads - similar to the Stanford-Binet items which involve copying a bead chain from memory, this subtest consists of 18 loose beads of three distinct shapes: cubical, spherical and cylindrical; and a set of 10 fixed patterns used by the examiner. There is a template which has two grooves, one for the model pattern and one used by the subject, who must construct from memory, the pattern which has been exposed for a set time. Patterns increase in length and complexity. A cover is used to conceal the pattern from the subject after the prescribed exposure.

3. Passalong - This subtest is similar to the sliding block puzzles which have appeared as popular children's games. The subject is required to slide tiles from a fixed starting position

in a frame to an arrangement which matches a mirror image pattern shown on a card. The task is clear and unambiguous. A card which is the same size as the frame is presented to the subject. Printed on it is the arrangement of tiles that the subject is required to match. The printed arrangement is the same size and color as the tiles (blue and red). The subject is required to reverse the arrangement of the tiles in the frame by sliding them so that they are in the same location as represented on the task card;

4. Form Assembly - This subtest consists of a single white plastic board 14" x 6" with seven set black pieces each of a different shape (square, triangle, etc.) glued to it. A grey plastic square 4" x 4" is on the right hand side of the form board. A black plastic stencil is placed on the grey square. The area of grey not covered by the black is identical to one of the seven designs on the form board. The items become progressively more difficult such that in later items the missing area requires two or three pieces to fill and thus completely covers the grey square;

5. Pattern Matching - This subtest is basically a derivative of the Kohs' Block Design Test (Kohs, 1923), also adopted by Wechsler. The main difference is that red and white square tiles are used instead of blocks, and the designs are constructed on a frame rather than in an unbound space as with the Block Design. Patterns are constructed from a card which is exposed for a set

time limit. The designs on the cards are the same size as the design to be constructed. There is a total of 15 items - some similar to items in various of the Wechsler Scales, some from the original Kohs' version, some new.

Scoring the test involves giving one point for each item correctly completed on each subtest. The total score is the sum of the five subtest scores. Total administration time is from forty-five minutes to one hour.

Concurrent and predictive validity were established (McElwain & Kearney, 1970) by demonstrating the relationship between test performance and measures of the criterion behavior, namely educational attainment, teachers' assessments of abilities, and performance on other well established tests of cognitive ability. In this particular case, the samples were 96 school children for the European sample and 241 school children for the Aboriginal sample (a medium contact sample). The tests chosen were the Otis Intermediate Test and the Raven Progressive Matrices Test. The Otis Test was found to be the most highly valid test for both the European and the Aboriginal samples. Correlations of test scores with the criterion behavior were positive for all the tests although significance failed to be reached at some age levels for the performance tests. This information refers only to one of the several studies and methods used in establishing validity.

Construct validity was first examined in the positive relationship between test scores and age, using the theoretical formulation 'If a set of scores derived from some instrument such as a test does not

show increases with age over the usual period of general biological maturation, then the scores are not measures of general cognitive capacity (McElwain & Kearney, 1970, p. 129)'. There was support for the hypothesis of a positive relationship between test scores and age in every sample. It is reported that the Rasch analysis and the factor studies of items indicate that the QT measures a general cognitive factor throughout its range.

Kearney (1969) reports the following results in her study based on a sample of 120 deaf children compared with a similar number of hearing children in the same grade and age range:

1. Raw score with age for both samples for each subtest showed a mean difference in test scores between the normal and deaf group for each age level with smaller mean differences for the older subjects (12 years) on the Form Assembly and Passalong subtests, and conversely, smaller mean differences for the younger subjects on the Pattern Matching and Knox Cube Test. There was no significant difference in the slope of the regression lines for raw score on age for both samples ($t = .196$). The difference term was also examined and at a t value of 6.78, indicated a significant difference between the two samples ($p < .001$).
2. Comparison of test-retest scores and of the Queensland Test with other IQ tests as discussed in Chapter II, p.p. 14 & 15.

Coloured Progressive Matrices

The Coloured Progressive Matrices (Raven, 1947) is a non-verbal test of intelligence which was developed by Raven in England, and initially

published in 1947. It consists of three sets of 12 problems or 36 problems in all. The test was originally constructed to supplement the 1938 Standard Progressive Matrices. Specifically, the Coloured Matrices consists of 24 problems or Sets A and B of the Standard Progressive Matrices to which were added an additional set of 12 problems, Set Ab. These problems are intermediate in difficulty to 5 and 12 of Set A and 1 and 7 of Set B. Since the test is non-verbal it may be used with individuals who do not speak the English language, with the deaf, with aphasics and others who have verbal comprehension difficulties. Raven describes his test as a measure of "observation and clear thinking". The problems were arranged to assess the chief cognitive processes of which children below age 11 are capable. It indicates clearly whether a person is or is not capable of forming comparisons and reasoning by analogy; and if not, to what extent relative to other people, he is capable of organizing spatial perceptions into systematically related wholes (Raven, 1956). The Raven's Test is not purported to be a measure of general intelligence. However, correlations of .50-.66 are reported between the test and the Binet Scale. Test-retest reliability is around .90.

PROCEDURE

Population and Sample

The population consisted of all students (153) enrolled at the Alberta School for the Deaf, Edmonton. They ranged in age from 5 years, 2 months to 20 years, 1 month. The sample selected for the

study consisted of 94 students (54 boys and 40 girls) between the age of 7 years, 0 months and 15 years, 0 months as of April 1, 1974. In terms of intellectual ability, the sample included all students in the above-mentioned age range regardless of their level of functioning in the classroom. It was not necessary to exclude any student due to severe visual or physical impairment which would have prevented a valid performance on the tests administered.

Data Collection

The School provided the following information on each child in the sample: name, birthdate, level of instruction, residence status, and level of hearing.

For the purpose of this study, residence status was categorized into three groups:

1. Day students - live at home and attend school daily;
2. Resident students - going home every weekend or every other weekend;
3. Resident students - going home less frequently than every other weekend, and including students who go to foster homes on weekends, long weekends or academic holidays only.

Once again for the purpose of this study, hearing level was categorized into three levels and based upon audiological data using the average decibel loss in the better ear within the speech awareness threshold:

1. 69 and below decibel loss;
2. 70 - 84 decibel loss;
3. 85 and above decibel loss.

In cases where a hearing loss was reduced through the use of a hearing aid, and reported, the "aided" hearing level was used as the figure for this data.

Test Administration

The Queensland Test was administered individually to the subjects by the writer, who was qualified to administer individual diagnostic tests but had not previously tested or worked with deaf children. There were 42 subjects who had been tested with the Queensland Test in May and June, 1973, by a qualified examiner who had experience with deaf and 'hard of hearing' children. It was decided to use their scores in the norming data to avoid a possbile inflation of scores as a result of retesting after only one year hence and the possibility of a practice effect.

The Test was administered in accordance with the standard procedure in the Test Manual. Instructions were communicated through pantomime, time allotments were observed, and a minimal use of 'signs' and 'finger-spelling' was employed to establish rapport and reinforce the subjects, notwithstanding the fact that the examiner's repertoire of this new language was comprised of essential expressions learned in a one-hour 'crash' session and with much enthusiastic urging from the school staff. No difficulty was encountered in administering the test.

The 1947 Coloured Progressive Matrices Test was administered as a group test to individual classes. The largest group had 10 students. In each class, the teacher assisted in giving the instructions and proctoring the session. Items A1 and A2 of Set A were used for

demonstration purposes and the answers revealed to the class. When an individual arrived at the first item of Set Ab or Set B, he was cued, shown where to proceed on the answer sheet, but the correct responses were not given.

Scoring both the Queensland Test and the Progressive Matrices was done by hand and performed by the examiner. Only 50 scores on the Coloured Progressive Matrices Test were included in subsequent analysis. Forty-two cases were rejected on the basis of having been tested on the Queensland Test one year ago. Two subjects were absent when the test was administered.

To facilitate later handling, the data for each subject - name, sex, birthdate, residence status, hearing level, age when tested, and test results, were entered on a 4" x 5" card, and the cards put in order of chronological age.

DESIGN AND ANALYSIS

Processing of Data

1. Means and Standard Deviations were computed for each age level on both the Queensland Test and the Coloured Progressive Matrices, and the results reported in table form (see Table II and Appendix B).
2. Raw scores for both tests and appropriately coded scores for the variables of sex, residence status, and hearing level, were transferred to IBM cards for computer analysis, which was carried out through the Division of Educational Research Services, to obtain product-moment correlations and a three-way analysis of

variance, respectively. The data were processed on the IBM System/360, in the Computing Centre, University of Alberta.

3. Frequency distributions of the Queensland Test raw scores for each age level were developed and the medians were computed on a graph. The same procedure was followed in obtaining the 1st and 3rd quartiles for each age level.

4. A table of Quartile Norms was developed after necessary smoothing of the norms at certain age levels was carried out.

Analysis

1. a product-moment correlation was used to test the hypothesis that: the Queensland Test will have a significant positive correlation with the Coloured Progressive Matrices, which was selected as a suitable test of non-verbal cognitive ability from which a comparison could be drawn.

2. A three-way analysis of variance was used to test the null hypotheses that:

a. there is no significant difference in the performance of boys and girls on the Queensland Test;

b. there is no significant difference in the performance of students of different levels of hearing on the Queensland Test;

c. there is no significant difference in the performance on the Queensland Test of students of residence status categories 1, 2 and 3.

A further look at whether there is any interaction effect

between the sex, residence status and hearing level of the students was made.

Summary

A description of the instruments used in this study was presented. The population and sample were described. The procedure used in collecting the data, administering the tests and processing the data was explained. Finally, the procedure in setting up the norms was described and the hypotheses to be tested were stated.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

The purpose of this study was to develop preliminary norms for students at the Alberta School for the Deaf for the Queensland Test. The Queensland Test was then correlated with the Coloured Progressive Matrices using a product-moment correlation. A test was used to obtain a .05 level of significance which was deemed necessary for the rejection of the null hypothesis. The interaction of the variables of sex, residence status and hearing level with students' performance on the Queensland Test was analyzed using a three-way analysis of variance. A .05 level of significance was deemed necessary to reject the null hypotheses.

Norms

In setting up the norms, the small number of cases in a given age group precluded the use of standard score norms, such as stanines, or full percentile norms. Quartile norms were selected as being the most appropriate for this sample and they are reported in Table I.

As illustrated in Figures 1 and 2, the medians were somewhat erratic, as was the spread, especially at ages 9 and 13. It was therefore necessary to fit the medians to a smoothed curved line in order to provide practical norms for which the test scores would increase with age. A regression toward the means of each group, using a straight line equation (see Kearney, 1969, p. 7) was not deemed suitable for this data since the scores tend to level off at about age 13 and definitely at ages 14 and 15. This can be expected with scores on tests of mental ability.

TABLE I

QUEENSLAND TEST

QUARTILE NORMS*

ALBERTA SCHOOL FOR THE DEAF

		Chronological Age in Years							
Quartiles	Percentiles	7	8	9	10	11	12	13	14-15
Q_3	75	30	33	37	41	44	47	48	49
MEDIAN	50	23	27	31	34	38	39	40	41
Q_1	25	10	15	20	25	30	33	34	35
N		10	10	14	15	12	8	15	10

* Smoothed norms as presented in Figures 1 and 2

FIGURE 1

QUEENSLAND TEST

MEDIAN SCORES FOR EACH AGE LEVEL

ALBERTA SCHOOL FOR THE DEAF

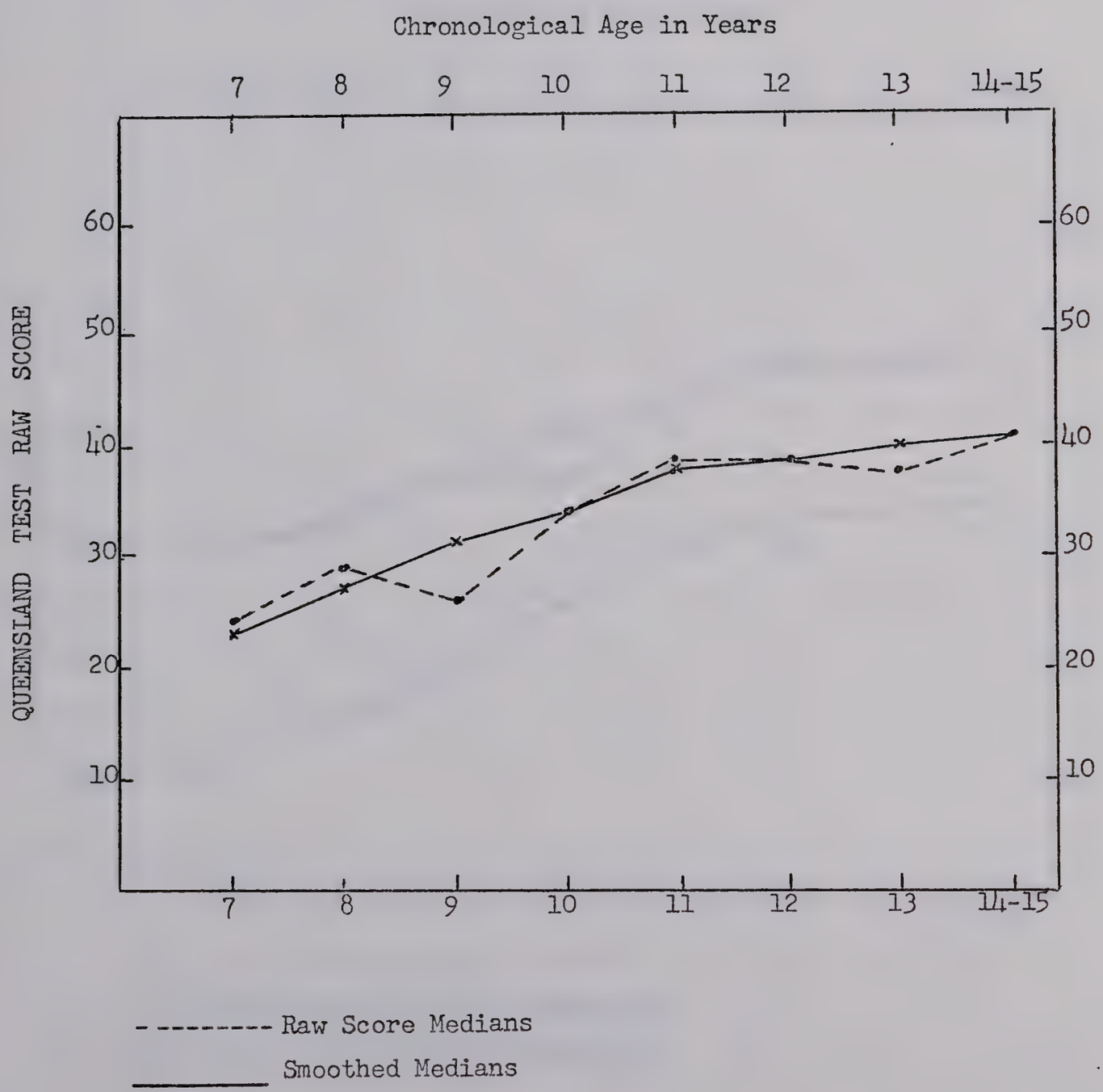
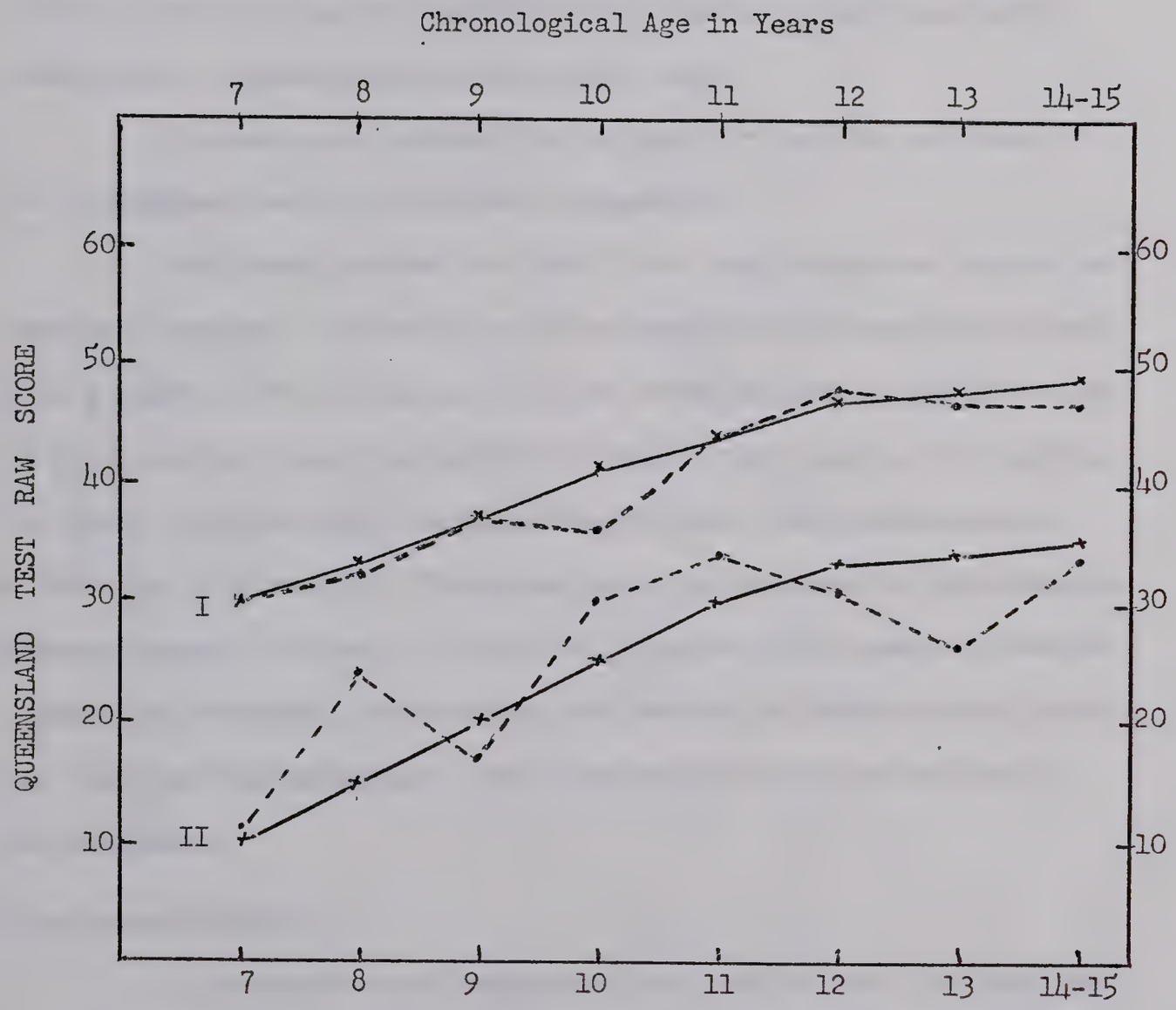


FIGURE 2

QUEENSLAND TEST

QUARTILE SCORES FOR EACH AGE LEVEL

ALBERTA SCHOOL FOR THE DEAF



I - Third Quartile (75th Percentile)

II - First Quartile (25th Percentile)

---------- Raw Score Quartiles

_____ Smoothed Quartiles

A small number of cases at ages 14 and 15 necessitated the combining of the two age groups for the purpose of presenting the norms.

Means and standard deviations for the Queensland Test are reported in Table II, which further illustrates the fluctuation in the norms as indicated in Figures 1 and 2. A larger sample would have the effect of stabilizing the distribution of scores at each age level, resulting in a smoother progression with age.

The means and standard deviations for the five subtests of the Queensland Test are reported in Appendix A.

The norms provided in Table I are preliminary and should be used with caution. For example, if we consider a 10 year old student with a score of 37, we can say that he scored higher than half of the 10 year olds but below the top 25% of the 10 year olds in the school. In short, his score falls between the 50th and 75th percentiles or within the 3rd quartile. This does not allow the user to discriminate between scores of 37 and 39 at age 10. However, the norms do provide a basis for comparing the relative performance of students at a given age level on the Queensland Test, when used at the Alberta School for the Deaf.

Hypotheses Testing

1. A product-moment correlation was used to test the hypothesis that: the Queensland Test will have a significant positive correlation with the Coloured Progressive Matrices Test. A significance level of .05 was deemed necessary to reject the null hypothesis.

TABLE II

QUEENSLAND TEST

SUMMARY OF MEANS AND STANDARD DEVIATIONS

ALBERTA SCHOOL FOR THE DEAF

Chronological Age in Years

	7	8	9	10	11	12	13	14-15
MEAN	22.4	30.2	27.5	32.53	39.5	38.0	35.266	41.3
S.D.	10.581	9.138	13.506	6.66	6.921	10.1	11.498	7.334
N	10	10	14	15	12	8	15	10

The results are reported in Table III. The obtained values were statistically significant to reject the null hypothesis.

TABLE III

CORRELATION OF QUEENSLAND TEST AND COLOURED PROGRESSIVE MATRICES FOR 50 DEAF STUDENTS, AGES 7-15

r	t	df	P
.80	9.27	48	.001

The Queensland Test compares favourably with the Coloured Progressive Matrices for the sample under study, and using the latter test as a criterion, the results indicate the Queensland Test is measuring non-verbal cognitive ability which shows progress with an increase in age up to a certain level of maturation.

2. A three-way analysis of variance was used to test the null hypotheses that:

a. there is no significant difference in the performance of boys and girls on the Queensland Test;

b. there is no significant difference in the performance of students of different levels of hearing on the Queensland Test;

c. there is no significant difference in the performance of students of different residence categories on the Queensland Test.

A further look at whether there was any interaction effect between the sex, residence status, and hearing level of the students was also made. A significance level of .05 was deemed necessary to reject the null hypothesis in each case. The results are reported in Tables IV and V.

The results supported the null hypothesis in each case and further analysis was not necessary.

The results indicate that a deaf student's frequency of contact with the hearing community, such as his family, does not have a direct effect on his performance on the Queensland Test. Further, a deaf student's relative level of hearing has no direct effect on his performance on the Queensland Test. Finally, boys and girls can be expected to perform comparatively on the Queensland Test. There was no indication of an interaction effect between the variables under investigation.

One must be careful in arrving at a firm conclusion from the results of this study. The analysis of variance was carried out with relatively low numbers of subjects in the appropriate cells. Perhaps if the sample population was greatly increased, the findings would have differed.

A relatively high correlation was shown to exist between the Queensland Test and the Coloured Progressive Matrices. This correlation, in part, is due to relationships other than that which is strictly classified as cognitive ability.

TABLE IV

QUEENSLAND TEST

CELL MEANS FOR THREE VARIABLES

RESIDENCE STATUS		HEARING LEVEL 1	2	3
1	M	40.87	36.42	25.0
	F	28.66	33.37	36.66
2	M	36.17	34.20	33.25
	F	32.60	24.5	30.75
3	M	30.17	29.0	37.66
	F	40.0	29.0	31.75

TABLE V

SUMMARY OF ANALYSIS OF VARIANCE

FOR DATA OF TABLE IV

Source	SS	df	MS	F	P
Sex (A)	56.48	1	56.48	.402	.527
Residence Status (B)	36.66	2	18.33	.131	.877
Hearing Level (C)	186.14	2	93.07	.663	.318
A X B	135.07	2	67.53	.481	.619
A X C	93.88	2	46.94	.334	.716
B X C	268.12	4	67.03	.477	.751
A X B X C	748.61	4	187.15	.133	.265

General Observations

The means and standard deviations for 50 students on the Coloured Progressive Matrices, are reported in Appendix B. In addition, medians and quartile points were computed as presented in Appendix C, and compared with the standardized norms for hearing children as reported by Raven (1956, p. 38). It was necessary to combine certain age groups because of the small number of cases. Norms beyond age 11 were extrapolated as they are not reported beyond that age by Raven.

On the basis of a rough estimate, using medians and quartile points, the results indicate that the deaf sample compares favorably with the normative group of hearing children. This would not necessarily be the case, however, if the sample tested was compared to a similar sample of hearing children and tested for statistical significance.

Several students, age 12 and above scored very high on the Coloured Progressive Matrices and were subsequently administered Sets C, D and E of the Standard Progressive Matrices in order to investigate the limit of their performance on the test. In computing a total score, Set Ab was omitted as outlined in Raven (1956, p. 12). The mean and standard deviation for two groups: ages 12-13 and ages 14-15, are reported in Appendix D. This group comprised 19 students from the sample of 50 students used in the correlation study. Medians and quartile points are reported in Appendix E and compared with the normative group of hearing children as reported by Raven (1938, p. 12). The results indicate that the 19 deaf students performed within the normal range and compared favorably with hearing children on the Progressive Matrices, 1938.

It was not possible to obtain sufficient test-retest data to report the test-retest reliability of the Queensland Test for the normative group, or to report the standard error of measurement for the norms, as time did not allow the retesting of some 40 students who had been initially assessed with the Queensland Test in May and June of 1973. A small sample of 12 students, ages 8-14 were tested before the information on the students who had been previously tested was made available. A product-moment correlation was used to obtain a reliability coefficient and a t test, using a difference of means method (Ferguson, 1971, p.153), was used to determine the significance at the .05 level. The results are reported in Appendix F. A high test-retest reliability is indicated, and taking the value of the correlation coefficient squared, the results indicate that .824 or 82% of the variance of the retest scores is predictable from the variance of the scores of the first test (Ferguson, 1971, pp. 115-117).

As stated earlier in this Chapter, the small number of cases precluded the development of standard score norms. The use of stanine norms for this project would have provided a more precise method of interpreting an individual score from the norm table. The use of stanine norms is clearly depicted in the Mackenzie District Norming Project (MacArthur, 1965).

In order to illustrate the use of stanines with the data on the Queensland Test from this study, norms were developed for two groups of 15 subjects, and reported in Appendix G.

An observation worthwhile mentioning pertains to the administration of the Queensland Test and the Progressive Matrices to students at the Alberta School for the Deaf. As one having had no previous experience with deaf children, I found the Queensland Test easy to administer according to the standardized procedure. This can be attributed first, to the pantomime directions which are used, and secondly, to the tasks which appeared to be of high interest to the subjects and relatively free of background distractions which might have confused the intent of the task. Timing on some of the items utilizes maximum limits and did not seem to evoke pressure on the subjects. Tasks such as the Passalong and Pattern Matching present a good basis for clinical observation of test behavior and problem-solving approach.

The Coloured Progressive Matrices did not present any undue problems when administered as a group test. It was noted, however, that several children demonstrated more frustration in working on this test as compared to the Queensland Test. It was difficult to determine whether this was caused by difficulty with the test items or with the method of responding on the separate answer sheet. Although the test was administered in a group situation, it was possible for the examiner and the classroom teacher to provide sufficient individual attention to all children thereby preventing an invalid performance caused by confusion with the test procedure.

Summary of Results

The development of norms which are reported as Quartiles, was presented, along with pertinent information on the Queensland Test. A correlation between the Queensland Test and the Coloured Progressive Matrices revealed a significant positive relationship ($r = .80$). The results from a three-way analysis of variance which was used to investigate the effect of sex, residence status and hearing level on a student's Queensland Test score were found to support the null hypotheses that there was no significant difference.

Additional observation revealed that the deaf students performed within the normal range on the Coloured and Standard Progressive Matrices.

CHAPTER V

SUMMARY AND RECOMMENDATIONS

Summary

The primary purpose of this thesis was to establish preliminary norms for students at the Alberta School for the Deaf, for the Queensland Test. The Test was selected as an instrument which shows promise in assessing the non-verbal intelligence or cognitive ability of deaf children. A comparison of the Queensland Test with the Coloured Progressive Matrices was made, using a product-moment correlation. The results ($r = .80$) indicated a significant relationship between the two tests. The relationship between the effects of sex, residence status, and hearing level, and a student's performance on the Queensland Test were investigated using analysis of variance. The results indicated no significant relationship. In view of the small numbers used in the analysis no firm conclusions should be made. Quartile norms were established. The small number of cases at each age level precluded the establishment of full percentile or standard score norms, such as stanines.

Recommendations

1. The Queensland Test can be used readily with deaf children as a performance scale of cognitive ability. However, results should be treated with caution when using the preliminary norms which have been developed, and other test results should be used in conjunction with them.

2. Norms should be built up within the next few years by adding to the sample tested in this study, possibly by obtaining data from other schools for the deaf which may be using the test, or by

conducting testing programs with the Queensland Test at the schools for the deaf in our neighboring provinces of British Columbia and Saskatchewan. Such a project would make possible the development of standard score norms.

3. Using the current sample, a "retest" should be conducted in one year's time in order that the test-retest reliability and standard error of measurement can be established.

4. The Queensland Test should be validated further using the WISC (Performance) or the Hiskey-Nebraska (Performance) as a criterion, since the literature reviewed indicated conflicting results on the use of the Progressive Matrices with the deaf. Also, the Progressive Matrices was administered as a group test and it would be more valid to make a comparison with an individual performance scale.

5. With respect to the use of the Queensland Test as a clinical tool, research should be directed toward an analysis of the subtests to provide such information as intercorrelations, factors being measured by each subtest, and more specific application of results to the needs of those being tested, such as the deaf.

6. A representative sample of normal hearing population should be tested with the Queensland Test in order to provide comparative norms.

7. With respect to the basis on which the Queensland Test was first developed, namely for pre-literates or in situations of reduced communication, the usefulness of this test in the assessment of native people in Alberta or in the North needs to be established.

SELECTED REFERENCES

American Psychological Association, Standards for Educational and Psychological Tests and Manuals. Washington, 1966.

Bean, K. L. Construction of Educational and Personnel Tests. New York: McGraw-Hill, 1953.

Birch, H. G. Introduction in Developmental Potential of Preschool Children, by Else Haeussermann. New York: Grune & Stratton, Inc., 1958.

Birch, Jane and Birch, J.W., The Leiter International Performance Scale as an aid in the psychological study of deaf children. American Annals of the Deaf, 1951, 96, 502-509.

Bonham, S. J., Predicting achievement for deaf children. Unpublished research, Dayton Public Schools and Division of Special Education, Ohio Dept. of Education, 1963.

Brown, F. G. Principles of Educational and Psychological Testing. Dryden Press, 1970.

Cartan, Sue, Use of the Queensland Test in the subnormal population. Australian Journal of Mental Retardation, 1971, 1 (8), 231-234.

Day, H., Fusfeld, I., and Pintner, R., A Survey of American Schools for the Deaf. Washington, D.C.: The National Research Council, 1928.

Drever, J., and Collins, M. Performance Test of Intelligence. London, England, Oliver and Boyd, 1936.

Evans, L., A comparative study of the Wechsler Intelligence Scale for Children (Performance) and Raven's Progressive Matrices with deaf children. Teacher of the Deaf, March, 1966, 65, 76-82.

Ferguson, G. A. Statistical Analysis in Psychology and Education. McGraw-Hill, 1971.

Furth, H. G. Deafness and intellectual development, in Godfrey, B. (Ed.). Orientation of Social Workers to the Problem of Deaf Persons. Washington, D. C.: U. S. Department of Health, Education and Welfare, Vocational Rehabilitation Administration, 1964.

Furth, H. G. Deafness and Learning: A Psychosocial Approach. Belmont, Calif.: Wadsworth Publishing Company, Inc., 1973.

Furth, H. G. Thinking Without Language. New York: The Free Press, 1966.

Gaskill, P., Longworth, A., and Neilson, R. The administration of Snijders-Oomen Non-verbal Intelligence Tests to deaf children. Teacher of the Deaf, 1970, 68(400), 137-142.

Goetzinger, C. P., Wills, R. C., and Dekker, L. C. Non-language I.Q. tests used with deaf pupils. Volta Review, 1967, 69, 500-506.

Goetzinger, M. R. and Houchins, R. R. The 1947 Colored Raven's Progressive Matrices with deaf and hearing students. American Annals of the Deaf, 1969, 114, 95-101.

Goetzinger, C. P. and Rousey, C. L. A study of the Wechsler Performance Scale (Form II) and the Knox Cube Test with deaf adolescents. American Annals of the Deaf, 1957, 102, 388-398.

Graham, E. Ellis and Shapiro, Esther. Use of the Performance Scale of the Wechsler Intelligence Scale for Children with the deaf child. Journal of Consulting Psychology, 1953, 17, 396-398.

Guilford, J. P. A revised structure of intellect. Reports from the Psychological Laboratory of the University of Southern California, No. 19., 1957.

Guilford, J. P. and Frucnter, B. Fundamental Statistics in Psychology and Education. McGraw-Hill, 1973.

Hanson, Grace, Hancock, Betty, and Kopra, Lennart. Relationship among audiological status, linguistic skills, visual-motor perception, and academic achievement of deaf children. Unpublished research. University of Texas, Austin, July, 1969.

Hiskey, Marshall S. A study of the intelligence of deaf and hearing children. American Annals of the Deaf, 1956, 101 (4), 329-339.

Hiskey, Marshall S. Manual: Hiskey-Nebraska Test of Learning Aptitude. Lincoln (Nebraska): Union College Press, 1966.

Hiskey, Marshall S. Nebraska Test of Learning Aptitude. (Rev.) Lincoln: University of Nebraska, 1955.

Ives, L. A. Proposals for a British learning aptitude scale standardized on deaf, partially hearing and hearing children. Teacher of the Deaf, 1968, 66, 171-184.

Johnson, E. W. Let's look at the child and not the audiogram. Volta Review, 1967, 69, 306-310.

Kearney, Jacqueline E. A new performance scale of cognitive ability for use with deaf subjects. American Annals of the Deaf, 1969, 114, 2-14.

Kirk, S. A. and Weiner, B. (Eds.) Behavioral Research on Exceptional Children. Washington, D.C.: Council for Exceptional Children, NEA, 1963.

Klonoff, Harry. I.Q. constancy and age. Perceptual and Motor Skills, 1972, 35, 527-534.

Levine, Edna S. Mental assessment of the deaf child. Volta Review, 1971, 73, 80-104.

Levine, Edna S. The Psychology of Deafness. New York: Columbia University Press, 1960.

Lyons, V. W., et al. Report of the 1931 Survey of the Illinois School for the Deaf. American Annals of the Deaf, 1933, 78, 157-175.

Lord, F. M. Test norms and sampling theory. Journal of Experimental Education, 1959, 27, 247-251.

MacArthur, R. S. Construct validity of three New Guinea Performance Scale subtests: Central Eskimos and Nsenga Africans. Unpublished paper presented at the 1st International Association for Cross-cultural Psychology Conference, Hong Kong, August 1971.

MacArthur, R. S. Mackenzie District Norming Project. Ottawa: Northern Administration Branch, Department of Northern Affairs and National Resources, 1965.

MacKane, K. A Comparison of the Intelligence of Deaf and Hearing Children. Teachers College Contribution to Education, No. 585. New York: Teachers College, Columbia University, 1932.

McElwain, D. W. and Kearney, G. E., Queensland Test Handbook. North Blackburn, Victoria: The Dominion Press, 1970.

McNemar, Q. Psychological Statistics. Wiley & Sons, 1969.

Myklebust, H. R. The Psychology of Deafness, Second ed. New York: Grune and Stratton, 1964.

Myklebust, H. R. and Burchard, E. M. L. A study of the effects of congenital and adventitious deafness on intelligence, personality, and social maturity of school children. Journal of Educational Psychology 1945, 34, 321.

Neuhaus, Mary. Modifications in the administration of the W.I.S.C. Performance subtests for children with profound hearing losses. Exceptional Children, 1967, 33 (8), 573-574.

Oleron, P. A study of the intelligence of the deaf. American Annals of the Deaf, 1950, 95, 179-195.

Oyer, Herbert J. Auditory Communication for the Hard of Hearing. Englewood Cliffs, N. J., Prentice-Hall, 1966.

Ord. I. G. Mental Tests for Pre-Literates. London: Ginn and Company, 1970.

Parker, J. E. A preliminary investigation into a new performance scale of cognitive capacity for use with deaf subjects. Unpublished M.Sc. Qual. Thesis, University of Queensland Library, 1965.

Pintner, R., and Patterson, D., Some conclusions from psychological tests of the deaf, Volta Review, 1918, 20, 10-13.

Quigley, Stephen P. The deaf and the hard of hearing. Review of Educational Research, 1967, 39 (1), 103-123.

Raven, J. C. Guide to Using the Coloured Progressive Matrices. Dumfries, Crichton Royal, 1947.

Raven, J. C. Guide to Using the Coloured Progressive Matrices. Dumfries, Crichton Royal, 1956.

Raven, J. C. Guide to Using Progressive Matrices (1938). Dumfries, Crichton Royal, 1938.

Reamer, J. C. Mental and Educational Measurements of the Deaf. Psychological Monographs, No. 132, Princeton, N.J.: The Psychological Review Co., 1921.

Rosenstein, Joseph. Cognitive abilities of deaf children. Journal of Speech and Hearing Research, 1960, 3, 108-119.

Scherer, Patricia. Visual learning processes in deaf children. Unpublished doctoral dissertation, Northwestern University, Chicago, 1968.

Schick, H. F. A performance test for deaf children of school age. Volta Review, 1934, 34, 657, cited in Myklebust, H. R. The Psychology of Deafness, Second ed. New York: Grune and Stratton, 1964.

Silverman, Toby R. Categorization behavior and achievement in deaf and hearing children. Unpublished research, U. S. Department of Health, Education and Welfare, January, 1967.

Snijders, J. Th. and Snijders-Oomen, N. Non-verbal Intelligence Tests for Deaf and Hearing Subjects. J. B. Wolters: Croningen, 1958.

Streng, A. and Kirk, S. A. The social competence of deaf and hard of hearing children in a public day school. American Annals of the Deaf, 1938, 83, 244.

Thorndike, R. L. and Hagen, Elizabeth. Measurement and Evaluation in Psychology and Education, Third ed. New York: John Wiley & Sons, 1969.

Thurstone, L. L. Primary Mental Abilities. University of Chicago Press, 1957.

Vernon, McCay and Brown, Donald W. A guide to psychological tests and testing procedures in the evaluation of deaf and hard of hearing children. Journal of Speech and Hearing Disorders, 1964, 29 (4), 414-423.

Vernon, P. E. Analysis of cognitive ability. British Medical Bulletin, 1971, 27 (3), 222-226.

Wechsler, D. Wechsler Intelligence Scale for Children - Manual. New York: Psychological Corporation, 1949.

Wechsler, D. The Measurement and Appraisal of Adult Intelligence (Fourth ed.). Williams and Wilkins, Baltimore, 1958.

Willis, Diane J., Wright, Logan and Wolfe, John. W.I.S.C. and Nebraska Performance of deaf and hearing children. Optometric Weekly, 1972, 63, 1206-1209.

Wrightstone, J. W., Aronow, Miriam S., and Moskowitz, Sue. Developing reading test norms for deaf children. Test Service Bulletin, No. 98. New York: Harcourt-Brace, 1962.

Zeckel, A. and Van der Kolk, J. A comparative intelligence test of groups of children born deaf and of good hearing. American Annals of the Deaf, 1939, 84, 114-123.

APPENDIX A

QUEENSLAND SUBTESTS - ALBERTA SCHOOL FOR THE DEAF

MEAN SCORES

Chronological Age in Years

	7	8	9	10	11	12	13	14-15
Knox Cube	3.9	7.5	5.0	7.6	7.9	6.7	6.2	7.5
Beads	1.5	3.1	2.6	3.3	4.8	4.0	3.6	4.3
Passalong	2.6	3.8	2.9	3.6	5.2	4.7	5.0	5.9
Form Assembly	7.7	7.1	7.4	8.5	8.7	10.4	6.9	9.3
Pattern Matching	7.3	7.7	9.6	9.4	12.1	12.1	12.3	14.3
Total - QT	22.4	30.2	27.5	32.5	39.5	38.0	35.3	41.3
N	10	10	14	15	12	8	15	10

APPENDIX A (CONTINUED)

QUEENSLAND SUBTESTS - ALBERTA SCHOOL FOR THE DEAF

STANDARD DEVIATIONS

Chronological Age in Years

	7	8	9	10	11	12	13	14-15
Knox Cube	2.92	2.17	7.19	3.02	2.31	3.11	3.23	2.63
Beads	1.43	2.28	5.16	1.34	2.42	1.93	2.64	2.75
Passalong	1.26	1.03	2.28	1.45	1.03	1.58	1.46	.99
Form Assembly	3.23	2.23	3.46	2.50	2.67	2.45	3.36	2.50
Pattern Matching	3.40	4.76	4.70	3.72	2.54	3.48	4.08	1.16
Total - QT	10.58	9.14	13.51	8.66	6.92	10.10	11.50	7.34
N	10	10	14	15	12	8	15	10

APPENDIX B

COLOURED PROGRESSIVE MATRICES (1947)

SUMMARY OF MEANS AND STANDARD DEVIATIONS

ALBERTA SCHOOL FOR THE DEAF

FOR 50 STUDENTS, AGES 7-15

	Chronological Age in Years				
	7-8	9-10	11-12	13	14-15
Mean	18.2	20.9	30.6	29.0	32.6
S. D.	6.44	7.51	5.29	4.32	3.94
N	9	12	7	12	10

APPENDIX C

COLOURED PROGRESSIVE MATRICES (1947)

MEDIANS AND QUARTILE POINTS

ALBERTA SCHOOL FOR THE DEAF

FOR 50 STUDENTS, AGES 7-15 COMPARED WITH NORMAL SAMPLE

	Chronological Age in Years				
Quartile	7-8	9-10	11-12	13[b]	14-15[b]
Q_3	24.25 (21.0)[a]	25.5 (28.0)	34.75 (33.0)	34.0 (35.0)	34.0 (36.0)
Mdn	17.0 (18.0)	22.5 (24.0)	32.0 (30.0)	31.0 (32.0)	33.5 (33.0)
Q_1	12.75 (16.0)	14.5 (22.0)	24.25 (26.0)	24.0 (28.0)	32.75 (29.0)
N	9	12	7	12	10

a - Figures in parentheses are comparative norms for hearing children (Raven, 1956, p. 38).

b - Norms for ages 11 and beyond have been extrapolated since they are not reported past age 11.

APPENDIX D

PROGRESSIVE MATRICES (1938)

MEANS AND STANDARD DEVIATIONS

ALBERTA SCHOOL FOR THE DEAF

FOR 19 STUDENTS, AGES 12-15

	Age	
	12-13	14-15
Mean	38.7	44.8
S. D.	10.67	4.46
N	10	9

APPENDIX E

PROGRESSIVE MATRICES (1938)

MEDIANS AND QUARTILE POINTS

ALBERTA SCHOOL FOR THE DEAF

FOR 19 STUDENTS, AGES 12-15 COMPARED WITH NORMAL SAMPLE

	Age	
Quartile	12-13	14-15
Q_3	47 (47)[a]	48 (48)
Mdn	43 (43)	45 (44)
Q_1	26 (35)	40 (38)
N	10	9

a - Figures in parentheses are comparative norms for hearing children (Raven, 1938, p. 12)

APPENDIX F

QUEENSLAND TEST

TEST-RETEST RELIABILITY

ALBERTA SCHOOL FOR THE DEAF

FOR 12 STUDENTS, AGES 8-14

	Test A (1973)	Test B (1974)	r	t	df	P
Mean	35.17	38.0	.91	2.365	11	.02
S. D.	10.34	9.74				

$\alpha = .05$

APPENDIX G

QUEENSLAND TEST

STANINE NORMS: AGES 10 AND 13

ALBERTA SCHOOL FOR THE DEAF

	AGE	
Stanines	10	13
9	46-60	48-60
8	40-45	43-47
7	37-40	41-42
6	35-36	40-41
5	33-34	38-40
4	32-33	29-37
3	21-31	25-28
2	16-20	15-24
1	0-15	0-14
N	15	15